

НЕЧЕТКАЯ НЕЙРОСЕТЕВАЯ КЛАСТЕРИЗАЦИЯ ИНТЕНСИВНОСТИ МС-ПОТОКОВ СОБЫТИЙ

А.Ту Нгуен, А.М. Кориков, А.Туан Нгуен
Томский политехнический университет
nguyenanhtu@tpu.ru

Введение

Искусственные нейронные сети (НС) технологии занимают лидирующие позиции среди современных информационных технологий. В [1] предлагаются новые модели нечетких НС (ННС), использующих нечеткие функции активации (ФА) нейронов ННС. В настоящей работе исследуется возможность применения этих моделей ННС для решения задач кластеризации параметров МС (Markov Chain)–потоков событий, широко применяемых для решения многих прикладных задач в транспортных системах, системах связи и компьютерных сетях.

Наборы данных для обучения ННС

Эксперименты по нейросетевой кластеризации МС-потоков событий выполнены на основе ННС, структура которой может быть представлена в виде многослойной однонаправленной сети. ННС состоит из входного слоя, одного скрытого слоя и выходного слоя. Скрытый слой состоит из 10 нейронов с нечеткими ФА, виды которых для нечетких нейронных сетей второго типа представлены в [1]. В этой же работе представлены четыре вида нечетких ФА, которые используются далее в настоящей работе.

Входными данными являются потоки, состоящие из 500 случайных процессов с параметром λ_i ($i = 1, 2, \dots, 10$). Данные каждого процесса получены из 1000 моментов времени. То есть, входные данные можно представить в виде матрицы размерности 5000×1000 . Выходные данные также являются матрицей-вектором размерности 5000×1 , состоящей из значений $1, 2, \dots, 10$ соответственно с параметрами $\lambda_1, \lambda_2, \dots, \lambda_{10}$.

Алгоритм кластеризации K-средних

Для кластеризации параметров случайных процессов предлагается использовать алгоритм кластеризации K-средних, который является простым повторяющимся алгоритмом кластеризации, разделяющий определенный набор данных на заданное пользователем число кластеров k [2]. В нашей работе начальные точки алгоритма K-средних генерируются случайно и количество классов равно десяти.

Описание ФА

Приведем описание ФА вида б [1] (рис.1).

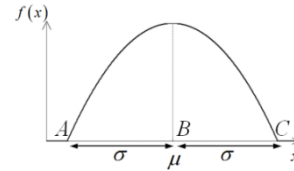


Рис. 1. Внешний вид ФА вида б

Набор точек $\langle A, B, C \rangle$ на рисунке 1 является характеристическими точками ФА.

Известно, что $f_L(x)$, $f_R(x)$ являются полиномами второго порядка и $f_L'(x) = 0$, $f_R'(x) = 0$ в характеристических точках, т.е.

$$f_L(x) = \sum_{i=0}^2 a_i x^i \quad f_R(x) = \sum_{i=0}^2 b_i x^i. \quad (1)$$

Для ФА вида б [1], коэффициенты полиномов $f_L(x)$ и $f_R(x)$ задаются следующими выражениями:

$$\begin{aligned} a_0 &= \frac{A^2}{(A-B)^2}; \quad a_1 = \frac{-2A}{(A-B)^2}; \quad a_2 = \frac{1}{(A-B)^2}; \\ b_0 &= \frac{C^2}{(B-C)^2}; \quad b_1 = \frac{-2C}{(B-C)^2}; \quad b_2 = \frac{1}{(B-C)^2}. \end{aligned} \quad (2)$$

Зададим $AB = BC = \sigma$; и $B = \mu$. Тогда $f_L(x)$ и $f_R(x)$ определяются следующим образом:

$$\begin{aligned} f_L(x) &= \frac{1}{\sigma^2} x^2 + \frac{-2(\mu-\sigma)}{\sigma^2} x + \frac{(\mu-\sigma)^2}{\sigma^2}, \quad \mu-\sigma \leq x \leq \mu \\ f_R(x) &= \frac{1}{\sigma^2} x^2 + \frac{-2(\mu+\sigma)}{\sigma^2} x + \frac{(\mu+\sigma)^2}{\sigma^2}, \quad \mu \leq x \leq (\mu+\sigma) \end{aligned} \quad (3)$$

μ и σ называются центром и шириной функций активации.

Пусть $a = (\mu - \sigma)$ если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$ если $\mu \leq x \leq (\mu + \sigma)$, тогда получим общее описание (обобщенная) ФА вида с:

$$f(x) = \frac{1}{\sigma^2} x^2 + \frac{-2a}{\sigma^2} x + \frac{a^2}{\sigma^2}. \quad (4)$$

Аналогично описываются ФА других видов [1].

Вычисление центров и ширины ядра функций активации

Алгоритм кластеризации K-средних используются для определения центра и ширины ФА нейронов ННС. Каждый кластер соответствует ядру ФА. Расположение центров ядра ФА определяется с помощью следующего уравнения:

$$\mu_i = \frac{1}{p} \sum_{j=1}^p X_j, \quad (5)$$

где μ_i является центром i -ого нечеткого скрытого нейрона; p - количество данных в формируемом кластере; X_j - j -ый данный в i -ом кластере.

Настройка ширины ядра является трудной проблемой при построении ННС. Ширина ядра устанавливается на среднее расстояние между данными соответствующего кластера:

$$\sigma_i = \sqrt{\frac{1}{p} \sum_{j=1}^p (X_j - \mu_i)^2} \quad (6)$$

где σ_i является шириной i -ого нечеткого скрытого нейрона, μ_i является центром i -ого нечеткого скрытого нейрона. Ширина ядра (5) гарантирует, что отдельные ФА не слишком острые или/и не слишком плоские.

Обучение ННС с использованием тестовых наборов данных

При построении ННС, задача корректировки весов сети между скрытым слоем и выходным слоем может быть отнесена к задачам оптимизации, обычно решаемыми методом наименьших квадратов. Веса ННС между скрытым слоем и выходным слоем рассчитываются с использованием метода псевдообращения [3].

X_1, X_2, \dots, X_N представляют собой входные наборы данных о пространственных входных признаках, которые представляют собой реализации случайных процессов. Выход y ННС в наблюдаемый момент времени j ($j = 1; 2; \dots; N$, N - количество наблюдаемых времени) имеет следующий вид

$$y_j = \sum_{i=1}^k w_i f(x_i), \quad j = 1; 2; \dots; N \quad (7)$$

где w_i - вес между выходным нейроном и i -ым скрытым нейроном, $f(x)$ - обобщенная ФА, k - число нейронов скрытого слоя.

Уравнение (6) можно записать в виде:

$$Y = \Phi w$$

$$\Phi(\mu, \sigma, X) = \begin{bmatrix} \varphi_1(\mu_1, \sigma_1, X_1) & \dots & \varphi_k(\mu_k, \sigma_k, X_1) \\ \dots & \dots & \dots \\ \varphi_1(\mu_1, \sigma_1, X_N) & \dots & \varphi_k(\mu_k, \sigma_k, X_N) \end{bmatrix}$$

$$v = 1; 2; \dots; N \quad (8)$$

φ - функция активации. Матрица Φ (размерность $N \times K$) называется выходной матрицей скрытого слоя нейронной сети; i -я строка Φ является i -м скрытым выходным нейроном относительно входов X_1, X_2, \dots, X_N .

Пусть T - матрица выходных данных обучения. Если $\Phi w = T$, то выходные веса w вычисляются по формуле

$$w = (\Phi^T \Phi)^{-1} \Phi^T T, \quad (9)$$

где $(\Phi^T \Phi)^{-1} \Phi^T$ является обобщенной псевдообратной выходной матрицей Мура-Пенроуза скрытого слоя.

Итак, основные параметры ННС определяются путем вычисления центров и ширины с использованием алгоритма кластеризации К-средних и определения весов выходного слоя ННС с использованием метода псевдообращения.

Результаты тестирования

Эксперименты проводились на базе процессора Core i5-4460 3,2 ГГц с 8 ГБ ОЗУ под MATLAB R2013a. В эксперименте набор данных тестирования генерируются случайно. Результаты эксперимента представлены на рисунке 2.

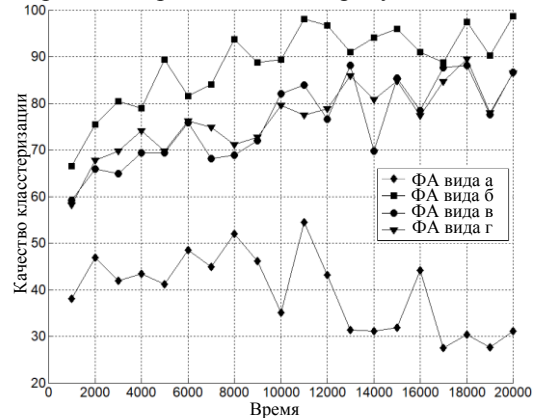


Рис. 2. Эффективность кластеризации по интенсивности МС-поток событий

Анализ графиков на рисунке 2 показывает что, разрабатываемые нами модели ННС успешно решают задачи кластеризации интенсивности МС-поток событий. При решении подобных задач ННС с ФА вида б имеют более высокую эффективность.

Заключение

Результаты исследования подтверждают эффективность применения ННС для решения задач кластеризации интенсивности марковских случайных потоков событий.

Список использованных источников

1. Нгуен А.Т., Кориков А.М. Модель нейронной сети с нечеткими функциями активации для прогнозирования временного ряда // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2016. – Т. 19. – № 4. – С. 50-52.
2. Пирсон Л. Наука данных для чайников / Л. Пирсон. – М.: Джон Уайли & сыновья, 2015. –
3. Сараев П.В. Использование псевдообращения в задачах обучения искусственных нейронных сетей // Электронный журнал "Исследовано в России". – 2001. - № 29. – С. 308-317.